

AperTO - Archivio Istituzionale Open Access dell'Università di Torino

## Discrete time Hammerstein model identification with unknown but bounded errors

### This is the author's manuscript

*Original Citation:*

*Availability:*

This version is available <http://hdl.handle.net/2318/6428> since

*Published version:*

DOI:10.1049/ip-cta:20010640

*Terms of use:*

Open Access

Anyone can freely access the full text of works made available as "Open Access". Works made available under a Creative Commons license can be used according to the terms and conditions of said license. Use of all other works requires consent of the right holder (author or publisher) if not exempted from copyright protection by the applicable law.

(Article begins on next page)

# Discrete time Hammerstein model identification with unknown but bounded errors

G.Belforte and P.Gay

**Abstract:** In this paper the problem of Hammerstein dynamic system identification is considered when nonlinear static blocks are described by generalised polynomials and linear time invariant blocks are modelled by ARX structures. The measurement error is characterised in a set membership context. The proposed approach accomplishes parameter identification introducing an extended Hammerstein model the parameter bounds of which can derive overbounds, which can, however, be tight up against to the Hammerstein model parameter uncertainties. The procedure for deriving such overbounds is presented in detail. The consistency of the algorithm for an increasing number of measurements is theoretically proved under the standard set membership assumption that theorises no overbounding of the measurement error, the extreme values of which always reoccur. The degree of conservativeness of the overbounds is evaluated through a simulation study based on a literature model and on a large set of randomly chosen systems. Both white noise and staircase inputs are considered. The results show that in most cases the derived overbounds are at most 10% larger than the actual bounds.

## 1 Introduction

Several real systems sometimes exhibit nonlinear behaviour that cannot be neglected or approximated by a proper linearisation technique, but must be somehow included in the model which describes the real system.

When this is the case, Hammerstein models are a possible option. They consist of a nonlinear static block followed by a linear dynamic block, as described in Fig. 1, where a discrete time version of such a model is reported. The inner signal  $x(k)$  is usually not available for measurement and only the system input  $u(k)$  can be chosen by the designer. The system output  $y(k)$  is measured with some additive measurement noise  $e(k)$ . Despite its relative simplicity, such a model is perceived to be quite adequate in describing real systems, such as chemical processes, distillation columns, electric heat exchangers, etc. (see e.g. [1] and [2]).

Identification of Hammerstein models has been studied since the late sixties, when Narendra and Gallman [3] proposed a two-step iterative identification procedure in which the parameters of the nonlinear block are fixed to an initial value, thus allowing the identification of the linear dynamic part and the evaluation of the inner signal  $x(k)$  from which the nonlinear block parameter estimates are derived and used as initial values for the next iteration.

This procedure may indeed not converge, as shown by Stoica [4].

In more recent years, the identification of Hammerstein models has been widely studied and several other identification algorithms have been proposed (see e.g. [3], [5–7]). While it is not intended to give a total overview here of the different identification algorithms established to date, we do notice that, to the best of our knowledge, in every case the measurement error  $e(k)$  has been assumed to be statistically described.

In this paper, the identification problem of Hammerstein dynamic systems is considered when the measurement error is characterised in a set membership context. The particular case of hard bounds on the measurements will be primarily considered.

The proposed approach accomplishes the parameter identification through the introduction of an extended Hammerstein model, the parameters of which are identified first using any algorithm available in the set membership literature. From the parameter bounds of such an extended model, overbounds to the Hammerstein model parameters are then derived. The main idea of this approach is similar to the one first proposed by Chang and Luus [8], although the latter only provides point estimates. The description of the measurement error in terms of membership sets allows us to derive guaranteed, although possibly conservative, estimates uncertainty bounds.

These overbounds converge to zero for increasing measurement number whenever the measurement error is not overbounded and its extreme values occur sufficiently

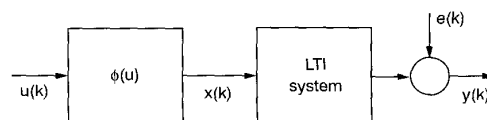


Fig. 1 Discrete-time SISO Hammerstein system

© IEE, 2001

IEE Proceedings online no. 20010640

DOI: 10.1049/ip-cta:20010640

Paper first received 8th December 2000 and in revised form 27th June 2001

G. Belforte is with the Dipartimento di Automatica e Informatica, Politecnico di Torino, Corso Duca degli Abruzzi, 24, 10129, Torino, Italy

P. Gay is with the Dipartimento di Economia e Ingegneria Agraria, Forestale e Ambientale, Università degli Studi di Torino, Via L. da Vinci, 44 10095 Grugliasco (To), Italy

often. Under such conditions, the proposed algorithm is consistent and allows one to arbitrarily derive low parameter uncertainties at the expense of longer data records. Moreover, simulation results show that the bounds derived with the proposed algorithms are in general not very conservative. Should the true parameter bounds be really needed, they can be evaluated using a Monte Carlo approach starting from the overbounds.

## 2 Problem formulation

Consider an SISO discrete-time Hammerstein model consisting of a static nonlinear block followed by a linear dynamic subsystem.

The static nonlinear part  $\phi(\cdot)$  is assumed to be a generalised polynomial so that the unmeasurable inner signal  $x(k)$  can be expressed as

$$x(k) = \phi(u(k)) = \sum_{i=1}^p c_i g_i(u(k)) \quad (1)$$

where  $g_1(\cdot), g_2(\cdot), \dots, g_p(\cdot)$  are given linear independent functions defined on the interval  $[u_a, u_b]$  with  $g_i(0)=0$ ,  $i=1, \dots, p$  so that the system steady state output is zero for zero input  $u$ .

The linear dynamic part is described by an ARX model

$$A(q)y(k) = B(q)x(k) + e(k) \quad (2)$$

with

$$B(q) = b_0 + b_1 q^{-1} + b_2 q^{-2} + \dots + b_{n_b} q^{-n_b}$$

$$A(q) = 1 + a_1 q^{-1} + a_2 q^{-2} + \dots + a_{n_a} q^{-n_a}$$

where  $q^{-1}$  is the usual delay operator.

The system input  $u(k)$  is bounded and can vary only on the given interval  $[u_a, u_b]$ , while the error  $e(k)$  is assumed to be unknown but bounded, so that

$$|e(k)| \leq E \quad \forall k \quad (3)$$

where  $E$  is a given constant.

Combining Relations (1) and (2) the following input/output relation can be derived

$$\begin{aligned} y(k) = & b_0[c_1 g_1(u(k)) + c_2 g_2(u(k)) + \dots + c_p g_p(u(k))] \\ & + b_1[c_1 g_1(u(k-1)) + c_2 g_2(u(k-1)) + \dots \\ & + c_p g_p(u(k-1))] \\ & \vdots \\ & + b_{n_b}[c_1 g_1(u(k-n_b)) + c_2 g_2(u(k-n_b)) + \dots \\ & + c_p g_p(u(k-n_b))] - a_1 y(k-1) \\ & - a_2 y(k-2) - \dots - a_{n_a} y(k-n_a) + e(k) \end{aligned} \quad (4)$$

This relation is bilinear in the parameters  $b$  and  $c$  while it is linear in the parameters  $a$ . Parameters  $b$  and  $c$  in Relation (4) are not uniquely related to the total system gain and cannot be uniquely identified. To avoid this ambiguity, assume, without loss of generality,  $b_0 = 1$ .

The  $a$ ,  $b$  and  $c$  parameters of Relation (4) can be rearranged, for notation convenience, into a single vector  $\xi$  as

$$\begin{aligned} \xi = & [\xi_1 \quad \xi_2 \quad \dots \quad \xi_{p+n_b+n_a}]^T \\ = & [c_1 \quad c_2 \quad \dots \quad c_p \quad b_1 \quad \dots \quad b_{n_b} \quad a_1 \quad \dots \quad a_{n_a}]^T \end{aligned}$$

With the set membership error assumption of Relation (3), the identification of the parameter vector  $\xi$  consists of

finding the admissible parameter set  $\mathcal{D}_\xi$  of all vectors  $\xi$  consistent with Model (4), the measurements  $y$  and the errors  $e$ . While any point in  $\mathcal{D}_\xi$  can be used as a parameter estimate, the 'size' of  $\mathcal{D}_\xi$ , usually measured in terms of its  $l_\infty$  diameter, is a measure of the estimate's reliability.

If interested in more details on the set membership identification, the reader can refer to [9–13] and the bibliography therein.

When a data set  $\{u(k), y(k)\}_{k=k_0}^{k=k_0+m-1}$  of  $m$  consecutive measurements is available, Relation (4) can be arranged as

$$y = F(\xi) + e$$

where  $y = [y(k_0) \ y(k_0+1) \ \dots \ y(k_0+m-1)]^T$  is the vector containing  $m$  consecutive measurements and  $e \in \mathbb{R}^m$  is the corresponding error vector while  $F(\xi)$  is the nonlinear function relating the parameters to the measurements.

The related admissible parameter set  $\mathcal{D}_\xi$  can be expressed as

$$\mathcal{D}_\xi = \{\xi \in \mathbb{R}^{p+n_b+n_a} : y = F(\xi) + e, e \in \Omega_e\} \quad (5)$$

where  $\Omega_e$  is the error membership set defined as

$$\Omega_e = \{e \in \mathbb{R}^m : |e(k)| \leq E, k = k_0, \dots, k_0+m-1\}$$

Let  $\xi_i^m$  and  $\xi_i^M$  be the extreme values of the parameter  $\xi_i$  in  $\mathcal{D}_\xi$

$$\xi_i^m = \min_{\xi \in \mathcal{D}_\xi} \xi_i \quad \text{and} \quad \xi_i^M = \max_{\xi \in \mathcal{D}_\xi} \xi_i$$

then the parameter uncertainty interval (PUI) of the  $i$ th parameter  $\xi_i$ , defined as

$$PUI_{\xi_i} = [\xi_i^m, \xi_i^M]$$

and its width

$$W(PUI_{\xi_i}) = \xi_i^M - \xi_i^m$$

can be introduced as a measure of the estimate reliability.

Let us now introduce the following augmented parameter vector  $\theta \in \mathbb{R}^n$ , with  $n = p(n_b+1) + n_a$

$$\theta = [\theta_1 \quad \theta_2 \quad \dots \quad \theta_n]^T$$

where

$$\begin{aligned} \theta_1 = c_1 \quad \theta_2 = c_2 \quad \dots \quad \theta_p = c_p \\ \theta_{p+1} = b_1 c_1 \quad \theta_{p+2} = b_1 c_2 \quad \dots \quad \theta_{2p} = b_1 c_p \\ \vdots \\ \theta_{pn_b+1} = b_{n_b} c_1 \quad \theta_{pn_b+2} = b_{n_b} c_2 \quad \dots \quad \theta_{p(n_b+1)} = b_{n_b} c_p \\ \theta_{p(n_b+1)+1} = a_1 \quad \theta_{p(n_b+1)+2} = a_2 \quad \dots \quad \theta_n = a_{n_a} \end{aligned} \quad (6)$$

This set of relations, compactly indicated as  $\theta = G(\xi)$ , allows us to rewrite Relation (4) as

$$\begin{aligned} y(k) = & \sum_{i=1}^p (\theta_i g_i(u(k)) + \theta_{p+i} g_i(u(k-1))) \\ & + \theta_{2p+i} g_i(u(k-2)) + \dots + \theta_{pn_b+i} g_i(u(k-n_b)) \\ & + \sum_{i=1}^{n_a} \theta_{p(n_b+1)+i} y(k-i) + e(k) \end{aligned} \quad (7)$$

In the following, this structure will be referred to as the extended Hammerstein model.

Note that the set of models described by Relation (7) is broader and includes the set of models described by

Relation (4). In fact, only if the  $\theta$  parameters of an extended Hammerstein model represented by Relation (7) satisfy the following relations

$$\begin{aligned} \frac{\theta_{p+1}}{\theta_1} &= \frac{\theta_{p-2}}{\theta_2} \quad \dots \quad = \frac{\theta_{2p}}{\theta_p} \\ \frac{\theta_{2p+1}}{\theta_1} &= \frac{\theta_{2p+2}}{\theta_2} \quad \dots \quad = \frac{\theta_{3p}}{\theta_p} \\ &\vdots \\ \frac{\theta_{pn_b+1}}{\theta_1} &= \frac{\theta_{pn_b+2}}{\theta_2} \quad \dots \quad = \frac{\theta_{p(n_b+1)}}{\theta_p} \end{aligned} \quad (8)$$

is it possible to derive the  $\xi$  parameters of the corresponding Hammerstein model.

Defining  $g(u(\cdot))$  as

$$g(u(\cdot)) = [g_1(u(\cdot)) \quad g_2(u(\cdot)) \quad \dots \quad g_p(u(\cdot))]^T$$

and  $y_{k-1}$  as

$$y_{k-1} = [y(k-1) \quad y(k-2) \quad \dots \quad y(k-n_a)]^T$$

Relation (7) can be rearranged as

$$y(k) = [g^T(u(k)) \quad g^T(u(k-1)) \quad \dots \quad g^T(u(k-n_b)) \quad y_{k-1}^T] \theta + e(k)$$

For notation compactness then, when  $m$  subsequent measurements are collected, Relation (7) can be expressed in vectorial form as

$$y = A(u)\theta + e \quad (9)$$

where

$$A(u) = \begin{bmatrix} g^T(u(k_0)) & g^T(u(k_0-1)) \\ g^T(u(k_0+1)) & g^T(u(k_0)) \\ \vdots & \vdots \\ g^T(u(k_0+m-1)) & g^T(u(k_0+m-2)) \\ \dots & g^T(u(k_0-n_b)) & y_{k_0-1}^T \\ \dots & g^T(u(k_0+1-n_b)) & y_{k_0}^T \\ & \vdots & \vdots \\ \dots & g^T(u(k_0+m-n_b-1)) & y_{k_0+m-2}^T \end{bmatrix} \quad (10)$$

is the regression matrix,

$$y = [y(k_0) \quad y(k_0+1) \quad \dots \quad y(k_0+m-1)]^T$$

is the vector containing  $m$  consecutive measurements and  $e \in \mathbb{R}^m$  is the corresponding error vector. Also in this case it is possible to introduce the admissible parameter set

$$\mathcal{D}_\theta = \{\theta \in \mathbb{R}^n: y = A(u)\theta + e, e \in \mathcal{E}_e\}$$

that, due to the linearity of Relation (9), is a convex polytope described by a subset of the following  $2m$  linear inequalities

$$y(k_0+k) - E \leq a_{k+1}^T \theta \leq y(k_0+k) + E \quad k = 0, \dots, m-1 \quad (11)$$

with  $a_k^T$  representing the  $k$ th row of  $A$ .

Assuming

$$\theta_i^m = \min_{\theta \in \mathcal{D}_\theta} \theta_i \quad \text{and} \quad \theta_i^M = \max_{\theta \in \mathcal{D}_\theta} \theta_i,$$

the parameter uncertainty intervals  $PUI_{\theta_i} = [\theta_i^m, \theta_i^M]$  and their amplitude  $W(PUI_{\theta_i}) = \theta_i^M - \theta_i^m$  can be introduced.

Finally, the set  $\hat{D}_\theta$  that is the image in  $\mathbb{R}^n$  of the set  $\mathcal{D}_\xi$  can be defined as

$$\hat{D}_\theta = \{\theta \in \mathbb{R}^n: \theta = G(\xi), \xi \in \mathcal{D}_\xi\}$$

recalling that  $\theta = G(\xi)$  is the compact form of Relation (6). Obviously  $\hat{D}_\theta \supseteq \mathcal{D}_\theta$ .

### 3 Identification of the Hammerstein model

In order to derive an estimate of the  $\xi$  parameter vector and to evaluate its reliability using the parameter uncertainty intervals  $PUI_{\xi_i}$ ,  $i = 1, \dots, p + n_b + n_a$ , it is convenient to operate first on the extended Hammerstein model deriving estimates and uncertainty intervals for its  $\theta$  parameters. The actual input-output identifiability of the extended Hammerstein model is, however, a prerequisite to such computation. Before analysing this identifiability, we note that the functions  $g_1(\cdot), \dots, g_p(\cdot)$ , describing the nonlinear block, have been assumed to be linear independent, so that, for any  $\alpha = [\alpha_1, \alpha_2, \dots, \alpha_p] \neq 0$ , their linear combination  $\alpha_1 g_1(u) + \alpha_2 g_2(u) + \dots + \alpha_p g_p(u)$  is not identically zero over the interval  $[u_a, u_b]$ .

Linear independence of  $g_1(\cdot), \dots, g_p(\cdot)$  is a necessary and sufficient condition for the existence of at least one set  $\mathcal{U} = \{u_1, \dots, u_p\}$  with  $u_i \in [u_a, u_b]$ ,  $i = 1, \dots, p$ , such that the matrix

$$G(\mathcal{U}) = \begin{bmatrix} g_1(u_1) & g_2(u_1) & \dots & g_p(u_1) \\ g_1(u_2) & g_2(u_2) & \dots & g_p(u_2) \\ \vdots & \vdots & & \vdots \\ g_1(u_p) & g_2(u_p) & \dots & g_p(u_p) \end{bmatrix} \quad (12)$$

is not singular [14]. Note that such non-singularity condition implies the input-output identifiability of the standing alone nonlinear block.

#### 3.1 Identifiability of the extended Hammerstein model

For the identifiability of the augmented parameter vector  $\theta$  the following theorem holds.

**Theorem 1:** The parameter vector  $\theta$  of the extended Hammerstein model is identifiable iff the functions  $g_1(\cdot) \dots g_p(\cdot)$  are linearly independent.

*Proof:*

**Necessary condition:** If the  $g_1(\cdot), \dots, g_p(\cdot)$  are not linearly independent then there exist two different parameter sets of the nonlinear block

$$\bar{c} = [\bar{c}_1 \quad \bar{c}_2 \quad \dots \quad \bar{c}_p]^T \quad \text{and} \quad \underline{c} = [\underline{c}_1 \quad \underline{c}_2 \quad \dots \quad \underline{c}_p]^T$$

such that

$$\sum_{i=1}^p \bar{c}_i^T g_i(u) = \sum_{i=1}^p \underline{c}_i^T g_i(u), \quad \forall u \in [u_a, u_b]$$

From this it follows that, for any input, the two parameter sets  $\bar{c}$  and  $\underline{c}$  generate the same output data, and so the system is not identifiable.

*Sufficient condition:* This can be proved by showing that the square matrix  $A$  obtained considering  $m=n$  suitable measurements is not identically singular. To this extent, consider the set of levels  $\mathcal{U} = \{u_1, \dots, u_p\}$  with  $u_i \in [u_a, u_b]$ , and the particular input sequence  $\tilde{u}$  constituted by a series of equispaced pulses that are  $n_b$  steps apart from each other and followed by  $n_a$  extra zero inputs

$$\tilde{u} = [\underbrace{0 \dots 0}_{n_b} u_1 \underbrace{0 \dots 0}_{n_b} u_2 \dots \underbrace{0 \dots 0}_{n_b} u_p \underbrace{0 \dots 0}_{n_a}] \quad (13)$$

with  $u_i \in \mathcal{U}$ ,  $i = 1, \dots, p$ .

The regression matrix in Relation (10), relative to the input sequence  $\tilde{u}$ , can be rewritten in the following block triangular form

$$\bar{A}(\tilde{u}) = \begin{bmatrix} A_{11} & A_{12} \\ 0 & A_{22} \end{bmatrix}$$

where

$$A_{11} = \begin{bmatrix} G(\mathcal{U}) & & 0 \\ & G(\mathcal{U}) & \\ & & \ddots \\ 0 & & G(\mathcal{U}) \end{bmatrix} \quad \text{and} \quad A_{22} = \begin{bmatrix} y_{k_0+p(n_b+1)-1} \\ y_{k_0+p(n_b+1)} \\ \vdots \\ y_{k_0+p(n_b+1)+n_a} \end{bmatrix} \quad (14)$$

The matrix  $A_{11}$  is block diagonal with the blocks defined in Relation (12) and is not identically singular because, when the functions  $g_1(\cdot), \dots, g_p(\cdot)$  are linearly independent, at least one set  $\mathcal{U} = \{u_1, \dots, u_p\}$  exists, so that the matrix  $G(\mathcal{U})$  is not singular.

The matrix  $A_{22}$  is the output regression matrix of the standing alone linear dynamic AR model. If  $g_1(\cdot), \dots, g_p(\cdot)$  are linearly independent then the input sequence of the standing alone AR subsystem  $\tilde{x} = \sum_{i=1}^p c_i g_i(\tilde{u})$  is not identically constant so that the AR subsystem is input/output identifiable, which implies the matrix  $A_{22}$  is not singular. Since  $A_{11}$  and  $A_{22}$  are not identically singular,  $\bar{A}(\tilde{u})$  is itself not singular and that implies the system's identifiability.  $\square$

From Theorem 1 it follows that the parameter vector  $\theta$  can be identified whenever  $m \geq n$  using standard algorithms that can be found in the literature (see e.g. [9–13, 15]). The amplitude of each parameter uncertainty interval  $W(PUI_{\theta_i})$  is a monotonic not-increasing function of the number of measurements  $m$ . This follows from the fact that any new measurement, adding information, can only reduce the admissible parameter set, thus making the estimate more reliable. If basic conditions about the error realisation  $e$  are satisfied [15–17], the amplitude of the parameter uncertainty intervals converges to zero for increasing  $m$ . Such conditions mainly require no over-bounding of the measurement error the extreme values of which must always reoccur. Then

$$\exists m_0 < \infty: \forall m \geq m_0 \quad W(PUI_{\theta_i}) = 0, \quad i = 1, \dots, n \quad (15)$$

Note that such a condition is not very strong since it is not possible to evaluate *a priori*  $m_0$ , which although finite, can be arbitrarily large.

The primary interest for the Hammerstein system identification relates indeed not to the  $\theta$  parameters but to the  $\xi$  parameters. These, however, are related to the  $\theta$  parameters so that their estimates and uncertainties can be derived from those of the  $\theta$  parameters according to the following considerations.

### 3.2 Derivation of $PUI_{\xi}$ for Hammerstein models

As previously noted, Hammerstein models are a subset of the extended Hammerstein models. Only for  $\theta$  vectors satisfying Relations (8) can a  $\xi$  vector representing the Hammerstein model of Relation (4) be computed. Remembering that the extended Hammerstein model has been constructed assuming

$$\begin{aligned} c_1 &= \theta_1 & c_2 &= \theta_2 & \dots & c_p &= \theta_p \\ a_1 &= \theta_{p(n_b+1)+1} & a_2 &= \theta_{p(n_b+1)+2} & \dots & a_{n_a} &= \theta_n \end{aligned}$$

as reported in Relation (6), it follows that outer bounds to the  $c$  and  $a$  parameters of the nonlinear block and of the autoregressive linear dynamic part are directly obtained from the  $PUI_{\theta_i}$ . Defining such outerbounds as  $\tilde{P}UI_{c_i}$  and  $\tilde{P}UI_{a_i}$ , we obtain

$$\tilde{P}UI_{c_i} = PUI_{\theta_i} \supseteq PUI_{c_i} \quad i = 1, \dots, p$$

$$\tilde{P}UI_{a_i} = PUI_{\theta_{p(n_b+1)+i}} \supseteq PUI_{a_i} \quad i = 1, \dots, n_a$$

As far as the  $b$  parameters of the moving average part are concerned, the situation is a little more complex. For them the following relations

$$\begin{aligned} b_1 &= \frac{\theta_{p+1}}{\theta_1} & b_1 &= \frac{\theta_{p+2}}{\theta_2} & \dots & b_1 &= \frac{\theta_{2p}}{\theta_p} \\ b_2 &= \frac{\theta_{2p+1}}{\theta_1} & b_2 &= \frac{\theta_{2p+2}}{\theta_2} & \dots & b_2 &= \frac{\theta_{3p}}{\theta_p} \\ & \vdots & & & & & \\ b_{n_b} &= \frac{\theta_{pn_b+1}}{\theta_1} & b_{n_b} &= \frac{\theta_{pn_b+2}}{\theta_2} & \dots & b_{n_b} &= \frac{\theta_{p(n_b+1)}}{\theta_p} \end{aligned} \quad (16)$$

should hold contemporaneously. This condition cannot be tested using the  $PUI_{\theta_i}$  only, although outer bounds  $\tilde{P}UI_{b_i} \supseteq PUI_{b_i}$  can be derived by computing

$$\tilde{P}UI_{b_i} = [\max\{\underline{b}_{i_1}, \underline{b}_{i_2}, \dots, \underline{b}_{i_p}\}, \min\{\bar{b}_{i_1}, \bar{b}_{i_2}, \dots, \bar{b}_{i_p}\}] \quad (17)$$

where

$$\underline{b}_{i_j} = \begin{cases} -\infty & \text{if } 0 \in [\theta_j^m, \theta_j^M] \\ \min \left\{ \frac{\theta_{ip+j}^m}{\theta_j^m}, \frac{\theta_{ip+j}^m}{\theta_j^M}, \frac{\theta_{ip+j}^M}{\theta_j^m}, \frac{\theta_{ip+j}^M}{\theta_j^M} \right\} & \text{otherwise} \end{cases} \quad (18)$$

and

$$\bar{b}_{i_j} = \begin{cases} +\infty & \text{if } 0 \in [\theta_j^m, \theta_j^M] \\ \max \left\{ \frac{\theta_{ip+j}^m}{\theta_j^m}, \frac{\theta_{ip+j}^m}{\theta_j^M}, \frac{\theta_{ip+j}^M}{\theta_j^m}, \frac{\theta_{ip+j}^M}{\theta_j^M} \right\} & \text{otherwise} \end{cases} \quad (19)$$

Note that the intersection of the intervals  $[\underline{b}_{i_j}, \bar{b}_{i_j}]$ ,  $j = 1, \dots, p$  cannot be empty as long as  $\mathcal{D}_{\xi}$  is not empty.

In fact, whenever a non empty  $\mathcal{D}_\xi$  set exists, its image  $\hat{D}_\theta$  ( $\hat{D}_\theta \supseteq \mathcal{D}_\theta$ ) is not empty either. For any point  $\theta^*$  in  $\hat{D}_\theta$  Relations (16) can be contemporarily solved by definition and lead to some  $\xi^*$ . Thus any component of  $b_i^*$  of  $\xi^*$  is related to the components  $\theta_i^*$  of  $\theta^*$  by

$$b_i^* = \frac{\theta_{ip+j}^*}{\theta_j^*}$$

Since  $\underline{b}_i$  and  $\bar{b}_i$  minimise and maximise  $\theta_{ip+j}/\theta_j$  whenever  $\theta_{ip+j}$  and  $\theta_j$  independently vary on intervals  $PUI_{\theta_{ip+j}}$  and  $PUI_{\theta_j}$  (that respectively include  $\theta_{ip+j}^*$  and  $\theta_j^*$ ) it follows that  $\underline{b}_i \leq b_i^* \leq \bar{b}_i$ ,  $\forall i$ .

Moreover, whenever the  $W(PUI_{\theta_i})$  shrink to zero so that  $\theta_i^m = \theta_i^M$  from relations (18–19) it follows that  $\underline{b}_i = \bar{b}_i$ .

From the convergence properties of the  $\theta$  parameters, shown in Relation (15), it follows that also the widths  $W(P\tilde{U}_{I_{a_i}})$ ,  $W(P\tilde{U}_{I_{b_i}})$  and  $W(P\tilde{U}_{I_{c_i}})$  of the outer bounds of the Hammerstein model parameters are monotonic not-increasing functions of the number  $m$  of used measurements and will contract when an increasing number of measurements is processed, eventually converging to zero.

#### 4 Simulation study

In order to obtain deeper insight into the performances of the algorithm, a first simulation study based on a Hammerstein model presented in the literature [3, 8, 18] was performed. The nonlinear block was a third degree polynomial of the form

$$x(k) = c_1 u(k) + c_2 u^2(k) + c_3 u^3(k) \quad (20)$$

The nominal parameter values of the ARX linear part and of the nonlinear block used for generating the data were

$$\begin{aligned} a &= [0.5 \quad -0.4 \quad -0.26 \quad -0.03] \quad b_0 = 1 \\ b &= [0.5 \quad 0.4 \quad 2] \quad c = [1 \quad 3 \quad 2] \end{aligned} \quad (21)$$

Indeed  $b_0$  does not need to be identified. Randomly generated uniformly distributed input strings of different lengths (300, 100, 50) were first applied to the system and randomly generated uniformly distributed measurement errors added to the outputs. In each trial, seven initial measurements, used to load the dynamic system, were discarded and from the remaining input–output pairs the parameter overbounds  $P\tilde{U}_{I_{\xi_i}}$ ,  $i = 1, \dots, 10$ , were computed using the proposed algorithm. From such overbounds, the true parameter bounds  $PUI_{\xi_i}$ ,  $i = 1, \dots, 10$ , were evaluated using a Monte Carlo approach. This consisted of randomly generating parameter vectors with all their components within the corresponding  $P\tilde{U}_{I_{\xi_i}}$ . For each vector, it was tested whether Relations (16) held. If they held, this implied that the vector belonged to the  $\mathcal{D}_\xi$  set.  $\xi_i^m$  and  $\xi_i^M$  were then approximated with the smallest and the largest value of  $\xi_i$  among the set of vectors for which Relations (16) held. This lead indeed to a conservative inner approximation of the  $PUI_{\xi_i}$ ,  $i = 1, \dots, 10$ .

Input strings were uniformly generated in the interval  $[-1, +1]$ . With such inputs the outputs turned out to have an average absolute value of about 4.5. Two different error levels were considered: one in the interval  $[-0.5, +0.5]$ , the second in the interval  $[-2.5, +2.5]$ , so that they were about  $\pm 10\%$  and  $\pm 50\%$  of the average absolute value of the signal.

500 different input and error strings were generated for each input length and error bound combination.

For each realisation the index

$$\eta = \min_i \frac{W(PUI_{\xi_i})}{W(P\tilde{U}_{I_{\xi_i}})} \quad i = 1, \dots, 10$$

was computed and used to evaluate the performance of the algorithm. Note that for large error realisations and short measurement records there are cases in which the index  $\eta$  is zero, implying the practical unidentifiability of the Hammerstein model.

The results are summarised in Table 1, where the minimum and the mean of index  $\eta$  over the 500 realisations are reported. Note that, in the computation of these values, cases in which  $\eta$  is zero were omitted. The number of such cases is also reported in the Table.

From these results it appears that, for the considered system, whenever the parameter vector is practically identifiable ( $\eta \neq 0$ ) the parameter overbounds  $P\tilde{U}_{I_{\xi_i}}$  are quite close to the true parameter bounds  $PUI_{\xi_i}$ . Obviously this happens to a larger extent when measurement strings are longer and measurement errors are lower.

In order to test whether the chosen system is particularly favourable with respect to the performances of the algorithm, a second simulation was performed in which 250 different parameter sets for the same system structure were randomly generated in the interval  $[-3, +3]$ . Only stable systems were considered. For each of them, 20 input and error strings of 300 samples were generated. The error bounds were fixed in each trial equal to  $\pm 10\%$  of the average absolute value of the output. Also in this case the index  $\eta$  for each realisation was computed.

For three parameter sets some of the 20  $\eta$  indexes were equal to zero. The results relative to these three parameter sets were discarded. The histogram of the  $247 \times 20 = 4940$   $\eta$  indexes relative to the 20 input and error realisations for each one of the 247 remaining parameter sets is reported in Fig. 2.

From these results it appears that, for the considered system structure, identification results are in most cases close to those obtained in the first simulation study. The parameter set used in the first simulation is therefore not particularly favourable.

A white noise input signal is the most common choice for system parameter identification purposes when no extra information on the system parameters is available, since it is a persistently exciting signal with components over the whole bandwidth. Indeed such a signal is not a typical input in normal system operation and there could be cases in which it could not even be used for identification purposes. To take this option into account and to test the possible influence of different input signals, a second simulation study was performed on the nominal model,

**Table 1: Minimum and average values of  $\eta$ , number of cases in which  $\eta=0$  for the different error levels and input string lengths (random inputs)**

Input length	Error $[-0.5, +0.5]$			Error $[-2.5, +2.5]$		
	min. $\eta$	mean $\eta$	Cases with $\eta=0$	min. $\eta$	mean $\eta$	Cases with $\eta=0$
50	0.6183	0.9114	0	0.4286	0.7644	432
100	0.6651	0.9593	0	0.4797	0.9208	32
300	0.9917	0.9999	0	0.9836	0.9996	0

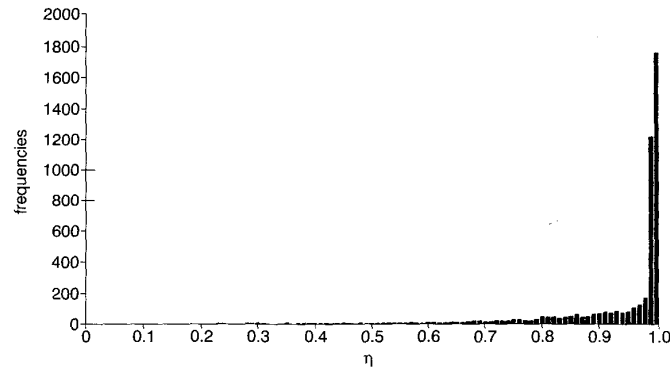


Fig. 2 Histogram of the 4940 different values of  $\eta$  obtained identifying 20 times each one of the 247 different randomly generated parameter sets

the parameters of which are reported in (21) using a different input signal. The number of simulation and error characteristics were exactly the same as in the previous study.

Since the given Hammerstein model is unidentifiable from step and square wave inputs because the parameters of the nonlinear block demand different input levels to be identified, the new chosen input was a staircase signal that also accurately approximated possible operating conditions in which different steady state reference levels are applied to the system.

For the 50 sample input strings, the staircase consisted of five series of 10 consecutive samples at the levels  $-1$ ;  $-0.5$ ;  $0$ ;  $0.5$ ;  $1$ . For the 100-sample input strings, the 50-sample input string was repeated twice while it was repeated six times for the 300-sample input string. The results relative to this simulation are reported in Table 2.

## 5 Conclusion

The identification of Hammerstein dynamic systems in which the nonlinear static block is described by a generalised polynomial and the linear time invariant block is modelled by an ARX model has been considered for the case of unknown but bounded measurement errors.

The proposed approach accomplishes the parameter identification through the introduction of an extended Hammerstein model, the parameters of which are first identified using any algorithm available in the set membership literature. From the parameter bounds of such an augmented model, overbounds to the Hammerstein model parameters are then derived.

Table 2: Minimum and average values of  $\eta$ , number of cases in which  $\eta=0$  for the different error levels and input string lengths (staircase inputs)

Input length	Error $[-0.5, +0.5]$			Error $[-2.5, +2.5]$		
	min. $\eta$	mean $\eta$	Cases with $\eta=0$	min. $\eta$	mean $\eta$	Cases with $\eta=0$
50	0.4506	0.8368	0	0.4011	0.6748	415
100	0.4243	0.8850	0	0.3418	0.7286	200
300	0.6218	0.9752	0	0.5501	0.9677	0

The consistency of the algorithm and thus its convergence to the 'true' parameter values for an increasing number of measurements has been addressed. It has been shown that the derived overbounds are monotonic not-increasing functions of the measurement number. Under standard assumptions in the set membership literature, which mainly require persistently exciting inputs and no measurement error overbounds so that the extreme error values always occur again, the parameter overbounds eventually shrink to a singleton that is the 'true' parameter vector. Therefore, with the proposed approach, arbitrarily low guaranteed estimate's uncertainties can be derived at the expense of longer data records.

The quantification of the conservativeness of the overbounds has been addressed in a simulation study. The obtained results show that the conservativeness is quite moderate.

## 6 Acknowledgments

This work was partially supported by Grant CNR 99.01960.ct07, MURST 40% MM07157922 and by IRITI-CNR, Torino.

## 7 References

- ESKINAT, E., JOHNSON, S.H., and LUYBEN, W.L.: 'Use of Hammerstein models in identification of nonlinear systems', *AIChE J.*, 1991, **37**, (2), pp. 255-268
- KORTMANN, M., and UNBEHAUEN, H.: 'Two algorithms for model structure determination of nonlinear dynamic systems with applications to industrial processes'. 8th IFAC Symposium on *Identification and system parameter estimation*, 1988, Beijing, China, pp. 939-946
- NARENDRA, K.S., and GALLMAN, P.G.: 'An iterative method for the identification of nonlinear systems using a Hammerstein model', *IEEE Trans. Autom. Control*, 1966, **11**, pp. 546-550
- STOICA, P.: 'On the convergence of an iterative algorithm used for Hammerstein system identification', *IEEE Trans. Autom. Control*, 1981, **26**, pp. 967-969
- AL-DUWAISH, H., and NAZMUL KARIM, M.: 'A new method for the identification of Hammerstein model', *Automatica*, 1997, **33**, (10), pp. 1871-1875
- HABER, R., and UNBEHAUEN, H.: 'Structure identification of nonlinear dynamic system—a survey on input/output approaches', *Automatica*, 1990, **26**, (4), pp. 651-677
- ZI-QIANG, L.: 'On identification of the controlled plants described by the Hammerstein system', *IEEE Trans. Autom. Control*, 1994, **39**, (3), pp. 569-573
- CHANG, F.H., and LUUS, R.: 'A non iterative method for identification using the Hammerstein model', *IEEE Trans. Autom. Control*, 1971, **6**, pp. 464-468
- 'Special issue on bounded-error estimation', *Int. J. Adapt. Control Signal Proc.*, 1994, **8**, (1)
- 'Special issue on bounded-error estimation (Part II)', *Int. J. Adapt. Control Signal Proc.*, 1995, **9**, (1)
- 'Special issue on parameter estimation with bounded error', *Math. Comput. Simul.*, 1990, **32**
- 'Special issue on trends in system identification', *Automatica*, 1995, **31**

- 13 'Special issue on system identification for robust control design', *IEEE Trans. Autom. Control*, 1992, **AC-37**, (7)
- 14 CHENEY, E.W.: 'Introduction to approximation theory' (McGraw-Hill, New York, 1966)
- 15 BELFORTE, G., and TAY, T.T.: 'Two new estimation algorithms for linear models with unknown but bounded measurement noise', *IEEE Trans. Autom. Control*, 1993, **38**, (8), pp. 1273–1279
- 16 BAI, E.W., CHO, H., and TEMPO, R.: 'Convergence properties of the membership set and noise model selection', *Automatica*, 1998, **34**, pp. 1245–1249
- 17 VERES, S.M., and NORTON, J.P.: 'Structure selection for bounded parameter models: consistency condition and selection criterion', *IEEE Trans. Autom. Control*, 1991, **36**, pp. 474–481
- 18 GALLMAN, P.G.: 'A comparison of two Hammerstein model identification algorithms', *IEEE Trans. Autom. Control*, 1976, **21**, pp. 124–126